Convergence to a Convention in a 2x2 Coordination Game with Adaptive Learning

Ethan Holdahl
University of Oregon
January 14, 2020

Coordination Game G

		Player B: 1	Player B: 0
Player A	1	$a_{11}, b_{11}{}^*$	a_{10}, b_{10}
	0	a_{01}, b_{01}	$a_{00}, b_{00}{}^*$

Lemma 1 will be used in the proof of Theorem 1.

Let G be a 2x2 coordination game, and let $P^{m,s,\epsilon}$ be adaptive learning with memory m, sample size s, and error rate ϵ. Denote the players A and B with actions $\{1, 0\}$ as depicted above. Define $h^t_{-i} = (x^{t-m}_{-i}, .., x^{t-1}_{-i})$ as the most recent m records of all players except player i at time t. Define R^t_i as the set of s records sampled by player i in period t from h^t_{-i}. Define BR^t_i as player i's best response to R^t_i.

Lemma 1. Let G be a 2x2 coordination game, and let $P^{m,s,\epsilon}$ be adaptive learning with memory m, sample size s, and error rate ϵ. If at any period n $\exists$ an action $x^* \in \{BR^n_i \cap BR^n_j\}$ for players i and j where $i \neq j$ then there exists a positive probability that each player plays x^* as a best reply for every period $e \geq n$.

Proof of Lemma 1. Assume at period n $\exists$ an action $x^* \in BR^n_A \cap BR^n_B$. I must show that there exists a positive probability that each player plays x^* as a best reply for every period $e \geq n$.

I use proof by induction.

Base Step: I show that if there exists an action x^* that is a best response for both players in period n then there exists a positive probability that both players play action x^* in period n.

Clearly, if x^* is a best response for both players then there exists a positive probability that both players play action x^* in period n.

Inductive Step: I show that $\forall e \geq n$ if action x^* was played as a best response for both players in period e then there exists a positive probability that action x^* is played as a best response in period $e+1$.

Assume action x^* was played as a best response for both players in period e. I must show that there exists a positive probability that action x^* is played as a best response in period $e+1$.

Note that since $x_i^t \in \{1, 0\} \quad \forall t$ the proportion of times that player $j \neq i$ played 1 in R_i^t, the set of s records in period t, is simply $\sum_{r \in R_i^t} \frac{r}{s}$.

Let $\alpha_i \in (0, 1)$ be the smallest probability that Player $j \neq i$ plays action 1 such that Player i's best response is playing action 1.

Without loss of generality assume that $x^* = 1$. That means action 1 is a best response for each player $i \in \{A, B\}$ to R_i^e, the set of s records sampled by player i in period e. That means that:

$$(1) \sum_{r \in R_i^e} \frac{r}{s} \geq \alpha_i$$

Now consider the set of s records sampled by player i in $e+1$: R_i^{e+1}.
In period $e+1$ each player i samples s records from h_{-i}^{e+1}. Note that $|h_{-i}^e \cap (h_{-i}^{e+1})'| = 1$, there is only 1 record in h_{-i}^e that is not in h_{-i}^{e+1}. Since $R_i^e \subseteq h_{-i}^e$ I know that $|R_i^e \cap (h_{-i}^{e+1})'| \leq 1$ That means there is at most 1 record in player i's sample in period e that is not able to be sampled in period $e+1$. This means that there is a positive probability that in period $e+1$ each player i samples $s-1$ records from the set $h_{-i}^{e+1} \cap R_i^e$ and the most recent record, x_{-i}^e. Assume both players samples in period $e+1$ fit this criteria and define $c_i^e = R_i^e \backslash R_i^{e+1}$, the record that was in the sample in period e but not in period $e+1$ for player i.
So I know that for each player i:

$$\frac{c_i^e}{s} + \sum_{r \in R_i^{e+1}} \frac{r}{s} = \frac{x_{-i}^e}{s} + \sum_{r \in R_i^e} \frac{r}{s}$$

However I know that $x_{-i}^e = 1$ and $c_i^e \in \{0, 1\}$. So $x_{-i}^e = 1 \geq c_i^e$.
Adding to both sides I get:

$$\frac{c_i^e}{s} + \frac{x_{-i}^e}{s} + \sum_{r \in R_i^{e+1}} \frac{r}{s} \geq \frac{x_{-i}^e}{s} + \frac{c_i^e}{s} + \sum_{r \in R_i^e} \frac{r}{s}$$

So

$$(2) \sum_{r \in R_i^{e+1}} \frac{r}{s} \geq \sum_{r \in R_i^e} \frac{r}{s}$$

Using (1) and (2) by transitivity I get:

$$\sum_{r \in R_i^{e+1}} \frac{r}{s} \geq \sum_{r \in R_i^e} \frac{r}{s} \geq \alpha_i$$

So for each player $i \in \{A, B\}$ action 1 is a best response to the sample R_i^{e+1}. Since action 1 is a best response for both players in period $e+1$ there exists a positive probability that each plays action 1 in period $e+1$.

Since both the base case and the inductive step has been shown, by mathematical induction I have shown that there exists a positive probability that each player plays x^* as a best reply for every period $e \geq n$. $\square$

Theorem 1. Let G be a 2x2 coordination game, and let $P^{m,s,\epsilon}$ be adaptive learning with memory m, sample size s, and error rate ϵ.

If $s < m$ then from any initial state, the unperturbed process $P^{m,s,0}$ converges with probability one to a convention and locks in.

Proof of Theorem 1. Define G as a 2x2 coordination game with adaptive learning where the possible actions for both players A and B are $\{1, 0\}$. Let memory $m \in \mathbb{N}$, sample size $s \in \mathbb{N}$ such that $s < m$, error rate $\epsilon = 0$ and let $h^t = (x^{t-m+1}, ..., x^t)$, be an arbitrary state at the end of period t. Let $\alpha \in (0, 1)$ be the smallest probability that Player B plays action 1 such that Player A's best response is playing action 1. Likewise, Let $\beta \in (0, 1)$ be the smallest probability that Player A plays action 1 such that Player B's best response is playing action 1.

There exists a positive probability that both players sample the most recent set of s records: $\{x^{t-s+1}, ..., x^t\}$ in period $t + 1$. Assume this is the case.

In period $t + 1$ the two players either
1) Share a best reply
or
2) Do not share a best reply
I will show a convention can be reached with positive probability in both cases.
Case 1: Both players share a best reply, x^*, in period $t + 1$. In this case I can apply Lemma 1 which shows that there exists a positive probability that both players play action x^* as a best reply for each period $e \geq t + 1$. If this happens then after period $e = t + m$ the entire memory is filled with both players playing action x^*. Since $\epsilon = 0$ and since both players could then only sample records of the other player playing x^* both would continue to play x^* as a best response for every period thereafter. So I have shown that there exists a positive probability that a convention can be reached and locked into with positive probability in Case 1.

Case 2: Assume the players do not share a best reply to the most recent set of s records. Since $\epsilon = 0$ they play different actions as best replies in period $t + 1$. Without loss of generality assume that in period $t + 1$ player A played action 1 and player B played action 0. This means that:

$$(3) \sum_{r=t-s+1}^{t} \frac{x_B^r}{s} > \alpha$$

and

$$(4) \sum_{r=t-s+1}^{t} \frac{x_A^r}{s} < \beta$$

Note: this is a strict inequality since the players do not share a best reply in period $t+1$ after sampling the set of records: $(x^{t-s+1}, ..., x^t)$.

Defining k and j

Assuming that player B continues to play action 0 for every period after period $t+1$. We know, since this is a coordination game, that if player A samples the most recent s actions in every period that there will exist a period, let's call the first one period $t+2+k$, where player A will have action 0 as a best response.

Thus, k is defined to be the smallest integer such that

$$(5) \sum_{r=t-s+2+k}^{t+1} \frac{x_B^r}{s} \leq \alpha$$

Note here that the record of x_B^r is 0 when $r = t+1$ and we assumed it is 0 for $r > t+1$. So the sum of the records x_B^r where $r > t+1$ is equal to 0 and drops out of this best response calculation. Since α is positive we know the inequality holds when $k = s-1$, and (3) tells us that it does not hold when $k = -1$. Thus, it is clear that for all histories and all α, $k \in \{0, ..., s-1\}$.

Likewise, let's assume that player A continues to play action 1 for every period after period $t+1$. We know, since this is a coordination game, that if player B samples the most recent s actions in every period that there will exist a period, let's call the first one period $t+2+j$, where player B will have action 1 as a best response.

Thus, j is defined to be the smallest integer such that

$$(6) \frac{j}{s} + \sum_{r=t-s+2+j}^{t+1} \frac{x_A^r}{s} \geq \beta$$

Note here that the record of x_A^r is 1 when $r = t+1$ and we assumed it is 1 for $r > t+1$. So the sum of the records $\frac{x_A^r}{s}$ where $r > t+1$ is equal to $\frac{j}{s}$.

Since β is less than 1 we know the inequality holds when $j = s - 1$, and (4) tells us that it does not hold when $j = -1$. Thus, it is clear that for all histories and all β, $j \in \{0, ..., s-1\}$.

Now I will prove that for all periods after $t+1$ and before $t+2+k$ player A's best response is to play action 1. Likewise, I will prove that for all periods after $t+1$ and before $t+2+j$ player B's best response is to play action 0.

Proving best responses for $e < k, j$

Define e as the amount of periods since $t+2$. When sampling the most recent set of s records in period $t+2+e$ player A has action 1 as a unique best response if:

$$(7) \qquad \sum_{r=t-s+2+e}^{t+1+e} \frac{x_B^r}{s} > \alpha$$

and player B has action 0 as a unique best response if:

$$(8) \qquad \sum_{r=t-s+2+e}^{t+1+e} \frac{x_A^r}{s} < \beta$$

Note that

$$(9) \qquad \sum_{r=t-s+2+e}^{t+1+e} \frac{x_i^r}{s} = \sum_{r=t-s+2+e}^{t+1} \frac{x_i^r}{s} + \sum_{r=t+2}^{t+1+e} \frac{x_i^r}{s}$$

So combining equations (7) and (9) where $i = B$, player A has action 1 as a unique best response if:

$$(10) \sum_{r=t-s+2+e}^{t+1} \frac{x_B^r}{s} + \sum_{r=t+2}^{t+1+e} \frac{x_B^r}{s} > \alpha$$

and combining equations (8) and (9) where $i = A$, player B has action 0 as a unique best response if:

$$(11) \sum_{r=t-s+2+e}^{t+1} \frac{x_A^r}{s} + \sum_{r=t+2}^{t+1+e} \frac{x_A^r}{s} < \beta$$

Consider $0 \leq e < k$. Since k is the smallest integer such that (5) $\sum_{r=t-s+2+k}^{t+1} \frac{x_B^r}{s} \leq \alpha$ it follows that $\forall e < k$, $\sum_{r=t-s+2+e}^{t+1} \frac{x_B^r}{s} > \alpha$.

Since $x_B^r \in [0, 1] \quad \forall r$, $\sum_{r=t+2}^{t+1+e} \frac{x_B^r}{s} \in [0, \frac{e}{s}]$.

So since $\sum_{r=t+2}^{t+1+e} \frac{x_B^r}{s} \geq 0$ and $\sum_{r=t-s+2+e}^{t+1} \frac{x_B^r}{s} > \alpha$ I get the condition (10):

$$\sum_{r=t-s+2+e}^{t+1} \frac{x_B^r}{s} + \sum_{r=t+2}^{t+1+e} \frac{x_B^r}{s} > \alpha$$

This condition means that for all integers e such that $0 \leq e < k$ when sampling the most recent s records in period $t+2+e$ that action 1 is a unique best response for player A.

Now consider $0 \leq e < j$. Since j is the smallest integer such that (6)$\frac{j}{s} + \sum_{r=t-s+2+j}^{t+1} \frac{x_A^r}{s} \geq \beta$ it follows that $\forall e < j$, $\frac{e}{s} + \sum_{r=t-s+2+e}^{t+1} \frac{x_A^r}{s} < \beta$.

Since $x_A^r \in [0,1] \quad \forall r$, $\sum_{r=t+2}^{t+1+e} \frac{x_A^r}{s} \in [0, \frac{e}{s}]$.

So since $\sum_{r=t+2}^{t+1+e} \frac{x_A^r}{s} \leq \frac{e}{s}$ and $\frac{e}{s} + \sum_{r=t-s+2+e}^{t+1} \frac{x_A^r}{s} < \beta$ I get the condition (11):

$$\sum_{r=t-s+2+e}^{t+1} \frac{x_A^r}{s} + \sum_{r=t+2}^{t+1+e} \frac{x_A^r}{s} < \beta$$

This condition means that for all integers e such that $0 \leq e < j$ when sampling the most recent s records in period $t+2+e$ that action 0 is a unique best response for player B.

There exists a positive probability that in each period $t+2+e$ where $0 \leq e < \min(k,j)$ both players can sample the most recent s records. I have just shown that when sampling the most recent s records in period $t+e$ where $0 \leq e < \min(k,j)$ player A best responds with action 1 and player B best responds with action 0. Assume both players do sample the most recent s records in periods $t+e$ where $0 \leq e < \min(k,j)$. Since $\epsilon = 0$ both players play their best response in those periods.

Proving coordination under different scenarios

I will now consider the three scenarios: $j < k$, $k < j$, and $j = k$.

First, $j < k$.

For period $t+2+j$ both players have a positive probability of sampling the most recent s records: $(x^{t-s+2+j}, ..., x^{t+1+j})$. Player B has a best response of action 1 in period $t+2+j$ if:

$$(12) \sum_{r=t-s+2+j}^{t+1} \frac{x_A^r}{s} + \sum_{r=t+2}^{t+1+j} \frac{x_A^r}{s} \geq \beta$$

Using (6) I know $\frac{j}{s} + \sum_{r=t-s+2+j}^{t+1} \frac{x_A^r}{s} \geq \beta$. And since $j < k$ I have already shown that player A has a unique best response of playing action 1 for all periods between $t+2$ and $t+2+j$ inclusive. So $\sum_{r=t+2}^{t+1+j} \frac{x_A^r}{s} = \frac{j}{s}$.
Which means (13) $\sum_{r=t-s+2+j}^{t+1} \frac{x_A^r}{s} + \sum_{r=t+2}^{t+1+j} \frac{x_A^r}{s} = \frac{j}{s} + \sum_{r=t-s+2+j}^{t+1} \frac{x_A^r}{s}$.
Combining (12) and (13) I get (6) which means that action 1 is a best response for Player B in period $t+2+j$.

Since both players can, with positive probability, sample the most recent s records and have the same action as a best response, I know by Case 1 a convention can be reached and locked into with positive probability when $j < k$.

Second, I consider $k < j$.

For period $t+2+k$ both players have a positive probability of sampling the most recent s records: $(x^{t-s+2+k}, ..., x^{t+1+k})$. Player A has a best response of action 0 in period $t+2+k$ if:

$$(14) \sum_{r=t-s+2+k}^{t+1} \frac{x_B^r}{s} + \sum_{r=t+2}^{t+1+k} \frac{x_B^r}{s} \leq \alpha$$

Using (5) I know $\sum_{r=t-s+2+k}^{t+1} \frac{x_B^r}{s} \leq \alpha$. And since $k < j$ I have already shown that player B has a unique best response of playing action 0 for all periods between $t+2$ and $t+2+k$ inclusive. So $\sum_{r=t+2}^{t+1+k} \frac{x_B^r}{s} = 0$.
Which means (15) $\sum_{r=t-s+2+k}^{t+1} \frac{x_B^r}{s} + \sum_{r=t+2}^{t+1+k} \frac{x_B^r}{s} = \sum_{r=t-s+2+k}^{t+1} \frac{x_B^r}{s}$.
Combining (14) and (15) I get (5) which means that action 0 is a best response for Player A in period $t+2+k$.

Since both players can, with positive probability, sample the most recent s records and have the same action as a best response, I know by Case 1 a convention can be reached and locked into with positive probability when $k < j$.

Third, I consider $k = j$.

In period $t + 2 + k$ player B can, with positive probability, sample the most recent set of s records. Since $j = k$ I have already shown that playing action 1 is a best response for player B in this scenario.

Since $m > s$ and both m and s are integers, I know that $m \geq s + 1$. So in period $t + 2 + k$ the records player A can sample from includes the most recent $s + 1$ records: $(x^{t+1+k-s}, ..., x^{t+1+k})$. So in period $t + 1 + k$ player A can, with positive probability, sample the set of s records: $(x^{t+1+k-s}, ..., x^{t+k})$. Note that these are the same records sampled in the previous period, $t + 2 + (k - 1)$, by player A which gave the unique best response of playing action 1.

So there exists a positive probability that both players share a best response in period $t + 2 + k$. So I can apply Lemma 1 here which shows that there exists a positive probability that both players play action 1 as a best reply for each period $e \geq t + 2 + k$. If this happens then after period $t + 1 + k + m$ the entire memory is filled with both players playing action 1. Since $\epsilon = 0$ and since both players could then only sample records of the other player playing action 1 both would continue to play 1 as a best response for every period thereafter. So I have shown that there exists a positive probability that a convention can be reached and locked into when $j = k$.

Thus, I have exhausted all three scenarios: $j < k$, $k < j$, and $j = k$ and shown that a convention can be reached with a positive probability in Case 2.

Since I have shown that a convention can be reached with positive probability in both Case 1 and Case 2 I have proven that from any initial state when $s < m$ and $\epsilon = 0$ a convention can be reached with positive probability and lock in. Since a convention can be reached from any arbitrary state and since conventions are absorbing states we know that as $T \to \infty$ that h^T, the state at time T converges with probability one to a convention in and locks in. ■

Forthcoming:
Theorem 2. Define G to be a 2x2 coordination game and let $P^{m,s,\epsilon}$ be adaptive learning with memory m, sample size s, and error rate ϵ. For all sufficiently large m (2?) the stochastically stable states of the perturbed process corresponds 1 to 1 with the risk dominant conventions.

Theorem 3. Define G to be a KxK coordination game and let $P^{m,s,\epsilon}$ be adaptive learning with memory m, sample size s, and error rate ϵ. Let (k,k) be Nash Equilibria for all $k \in \{1,...,K\}$. Define $i \in \{1,...,K\}$ and $j \neq i \in \{1,...,K\}$ to be conventions of both players playing action i in convention i and both players playing action j in convention j.
1. For all s, m the resistance from state i to state j is:

$$r(i,j) = \min(\lceil \alpha s \rceil, \lceil \beta s \rceil) + \max(\lceil \alpha s \rceil + \lceil \beta s \rceil - m, 0)$$

2. The risk dominance relation between state i and j remains unchanged for all $s/m \in (0,1]$ when m is sufficiently large.
3. In some, but not all, larger games (at least 3x3) which states are stochastically stable may depend on $s/m \in (0,1]$. If so, the value of $s/m \in (0,1]$ at which stochastic stability changes can be easily calculated.